



# UNITED STATES PATENT AND TRADEMARK OFFICE

UNITED STATES DEPARTMENT OF COMMERCE  
United States Patent and Trademark Office  
Address: COMMISSIONER FOR PATENTS  
P.O. Box 1450  
Alexandria, Virginia 22313-1450  
www.uspto.gov

APPLICATION NO.	FILING DATE	FIRST NAMED INVENTOR	ATTORNEY DOCKET NO.	CONFIRMATION NO.
09/629,831	07/31/2000	Jeffrey Thomas Kreulen	AM9-99-0157	4998
28211	7590	01/19/2006	EXAMINER	
FREDERICK W. GIBB, III GIBB INTELLECTUAL PROPERTY LAW FIRM, LLC 2568-A RIVA ROAD SUITE 304 ANNAPOLIS, MD 21401			SMITH, PETER J	
			ART UNIT	PAPER NUMBER
			2176	

DATE MAILED: 01/19/2006

Please find below and/or attached an Office communication concerning this application or proceeding.



UNITED STATES PATENT AND TRADEMARK OFFICE

---

Commissioner for Patents  
United States Patent and Trademark Office  
P.O. Box 1450  
Alexandria, VA 22313-1450  
[www.uspto.gov](http://www.uspto.gov)

**BEFORE THE BOARD OF PATENT APPEALS  
AND INTERFERENCES**

**MAILED**

**JAN 19 2006**

**Technology Center 2100**

Application Number: 09/629,831  
Filing Date: July 31, 2000  
Appellant(s): KREULEN ET AL.

Frederick W. Gibb, III  
For Appellant

**EXAMINER'S ANSWER**

This is in response to the appeal brief filed 11/8/2005 appealing from the Office action mailed 6/14/2005.

**(1) Real Party in Interest**

A statement identifying by name the real party in interest is contained in the brief.

**(2) Related Appeals and Interferences**

The examiner is not aware of any related appeals, interferences, or judicial proceedings which will directly affect or be directly affected by or have a bearing on the Board's decision in the pending appeal.

**(3) Status of Claims**

The statement of the status of claims contained in the brief is correct.

**(4) Status of Amendments After Final**

The appellant's statement of the status of amendments after final rejection contained in the brief is correct.

**(5) Summary of Claimed Subject Matter**

The summary of claimed subject matter contained in the brief is correct.

**(6) Grounds of Rejection to be Reviewed on Appeal**

The appellant's statement of the grounds of rejection to be reviewed on appeal is correct.

**(7) Claims Appendix**

The copy of the appealed claims contained in the Appendix to the brief is correct.

**(8) Evidence Relied Upon**

5,440,481	KOSTOFF et al.	8-1995
5,742,834	KOBAYASHI	4-1998
6,070,158	KIRSCH et al.	5-2000
6,470,307 B!	TURNER	10-2002

### **(9) Grounds of Rejection**

The following ground(s) of rejection are applicable to the appealed claims:

#### ***Claim Rejections - 35 USC § 103***

The following is a quotation of 35 U.S.C. 103(a) which forms the basis for all obviousness rejections set forth in this Office action:

(a) A patent may not be obtained though the invention is not identically disclosed or described as set forth in section 102 of this title, if the differences between the subject matter sought to be patented and the prior art are such that the subject matter as a whole would have been obvious at the time the invention was made to a person having ordinary skill in the art to which said subject matter pertains. Patentability shall not be negated by the manner in which the invention was made.

**Claims 1, 6 and 11 are rejected under 35 U.S.C. 103(a) as being unpatentable over Kostoff et al. (hereafter referred to as Kostoff), US 5,440,481 patented 08/08/1995.**

**Regarding independent claim 1**, Kostoff teaches determining a frequency of each word in each document in fig. 2, table 1, col. 4 lines 50-68, and col. 6 line 65 – col. 7 line 11. Kostoff teaches creating a table of most frequently occurring words in the documents in fig. 2, table 1, col. 4 lines 50-68, and col. 6 line 65 – col. 7 line 11. Kostoff teaches determining a frequency of phrases in each document that could contain only words in a table in fig. 2, table 1, col. 4 lines 50-68, and col. 6 line 65 – col. 7 line 11. Kostoff teaches outputting the most frequently occurring words and most frequently occurring phrases as a dictionary in fig. 2 and col. 4 lines 64-68.

Kostoff does not specifically teach inputting a maximum dictionary size and limiting the dictionary to the inputted maximum dictionary size, such that the dictionary contains less than all words in the documents. However, Kostoff does acknowledge the importance and limitation of

memory size for storing a list of trivial words in col. 4 lines 44-45. This list is a precursor to the dictionary, however it teaches one of ordinary skill in the art at the time of the invention the relevance of memory storage size. Kostoff also teaches selecting a portion of the word and phrase dictionary in col. 5 line 59 – col. 6 line 64. Kostoff uses an example of selecting the 60 most often repeated phrases. Kostoff notes that more or less than 60 most often repeated phrases may be selected at the discretion of the user.

In light of these teachings of Kostoff, one of ordinary skill in the art at the time of the invention would have truncated the dictionary of Kostoff at the user inputted number of most often repeated phrases in the event the dictionary had to reside within a limited memory storage. The teaching of Kostoff of possible memory storage constraints having an impact on a list size in col. 4 lines 44-45 would have motivated and taught insight to the person of ordinary skill in the art at the time of the invention to have made this modification. It would have been obvious to one of ordinary skill in the art at the time of the invention to have discarded the less frequent terms below the population threshold inputted by the user because they would not have been of further use in determining the themes of the text to prepare it for clustering with other documents. Eliminating the unused terms would have desirably saved memory as seen in col. 4 lines 44-45. Only the top set of words and phrases determined by the user would have been used and therefore it would have been obvious to have only retained those words and phrases in the dictionary.

**Regarding independent claim 6**, Kostoff teaches determining a frequency of each word in each document in fig. 2, table 1, col. 4 lines 50-68, and col. 6 line 65 – col. 7 line 11. Kostoff teaches creating a table of most frequently occurring words in the documents in fig. 2, table 1,

Art Unit: 2176

col. 4 lines 50-68, and col. 6 line 65 – col. 7 line 11. Kostoff teaches determining a frequency of phrases in each document that could contain only words in a table in fig. 2, table 1, col. 4 lines 50-68, and col. 6 line 65 – col. 7 line 11. Kostoff teaches outputting the most frequently occurring words and most frequently occurring phrases as a dictionary in fig. 2 and col. 4 lines 64-68.

Kostoff does not specifically teach inputting a maximum dictionary size and limiting the dictionary to the inputted maximum dictionary size, such that the dictionary contains less than all words in the documents. However, Kostoff does acknowledge the importance and limitation of memory size for storing a list of trivial words in col. 4 lines 44-45. This list is a precursor to the dictionary, however it teaches one of ordinary skill in the art at the time of the invention the relevance of memory storage size. Kostoff also teaches selecting a portion of the word and phrase dictionary in col. 5 line 59 – col. 6 line 64. Kostoff uses an example of selecting the 60 most often repeated phrases. Kostoff notes that more or less than 60 most often repeated phrases may be selected at the discretion of the user.

In light of these teachings of Kostoff, one of ordinary skill in the art at the time of the invention would have truncated the dictionary of Kostoff at the user inputted number of most often repeated phrases in the event the dictionary had to reside within a limited memory storage. The teaching of Kostoff of possible memory storage constraints having an impact on a list size in col. 4 lines 44-45 would have motivated and taught insight to the person of ordinary skill in the art at the time of the invention to have made this modification. It would have been obvious to one of ordinary skill in the art at the time of the invention to have discarded the less frequent terms below the population threshold inputted by the user because they would not have been of

Art Unit: 2176

further use in determining the themes of the text to prepare it for clustering with other documents. Eliminating the unused terms would have desirably saved memory as seen in col. 4 lines 44-45. Only the top set of words and phrases determined by the user would have been used and therefore it would have been obvious to have only retained those words and phrases in the dictionary.

Kostoff does not explicitly teach the creation of the word and phrases lists in two separate passes through the document. One of ordinary skill in the art at the time of the invention would have known how to create the two lists in separate passes through the document. It would have been obvious to one of ordinary skill in the art at the time the invention was made to use their skill in the art to have created each list as a result of each of two passes through the document. This would have been obvious and necessary in order to create the second list since the phrase selection would have been dependent on the contents of the first list.

**Regarding independent claim 11**, Kostoff teaches determining a frequency of each word in each document in fig. 2, table 1, col. 4 lines 50-68, and col. 6 line 65 – col. 7 line 11. Kostoff teaches creating a table of most frequently occurring words in the documents in fig. 2, table 1, col. 4 lines 50-68, and col. 6 line 65 – col. 7 line 11. Kostoff teaches determining a frequency of phrases in each document that could contain only words in a table in fig. 2, table 1, col. 4 lines 50-68, and col. 6 line 65 – col. 7 line 11. Kostoff teaches outputting the most frequently occurring words and most frequently occurring phrases as a dictionary in fig. 2 and col. 4 lines 64-68.

Kostoff does not specifically teach inputting a maximum dictionary size and limiting the dictionary to the inputted maximum dictionary size, such that the dictionary contains less than all

words in the documents. However, Kostoff does acknowledge the importance and limitation of memory size for storing a list of trivial words in col. 4 lines 44-45. This list is a precursor to the dictionary, however it teaches one of ordinary skill in the art at the time of the invention the relevance of memory storage size. Kostoff also teaches selecting a portion of the word and phrase dictionary in col. 5 line 59 – col. 6 line 64. Kostoff uses an example of selecting the 60 most often repeated phrases. Kostoff notes that more or less than 60 most often repeated phrases may be selected at the discretion of the user.

In light of these teachings of Kostoff, one of ordinary skill in the art at the time of the invention would have truncated the dictionary of Kostoff at the user inputted number of most often repeated phrases in the event the dictionary had to reside within a limited memory storage. The teaching of Kostoff of possible memory storage constraints having an impact on a list size in col. 4 lines 44-45 would have motivated and taught insight to the person of ordinary skill in the art at the time of the invention to have made this modification. It would have been obvious to one of ordinary skill in the art at the time of the invention to have discarded the less frequent terms below the population threshold inputted by the user because they would not have been of further use in determining the themes of the text to prepare it for clustering with other documents. Eliminating the unused terms would have desirably saved memory as seen in col. 4 lines 44-45. Only the top set of words and phrases determined by the user would have been used and therefore it would have been obvious to have only retained those words and phrases in the dictionary.



**Claims 2-5, 7-10, and 12-17 are rejected under 35 U.S.C. 103(a) as being unpatentable over Kostoff et al. (hereafter referred to as Kostoff), US 5,440,481 patented 08/08/1995 as applied to claims 1, 6, and 11 above, and further in view of Kirsch et al. (hereafter referred to as Kirsch), US 6,070,158 filed 11/13/1997, Kobayashi, US 5,742,834 patented 04/21/1998 and Turney, US 6,470,307 B1 filed 06/23/1997.**

**Regarding dependent claim 2,** Kostoff teaches adding words to a dictionary table in fig. 2, table 1, col. 4 lines 50-68, and col. 6 line 65 – col. 7 line 11. Kostoff teaches determining the frequency of each word remaining in the table in fig. 2, table 1, col. 4 lines 50-68, and col. 6 line 65 – col. 7 line 11. Kostoff teaches removing words below a frequency level from the dictionary table in col. 6 lines 2-64.

Kostoff does not teach removing punctuation and case from the documents. Kostoff does not teach removing stop words from the document. Kostoff does not teach replacing words in the documents with synonyms. Kostoff does not teach removing duplicate words from the documents. Kirsch teaches removing punctuation and case from the documents in col. 12 lines 5-7. Kirsch teaches removing stop words from the document in col. 12 lines 13-15. Kobayashi teaches replacing words in the documents with synonyms in fig. 3, 34-35, and col. 1 line 54 – col. 2 line 13. Turney teaches removing duplicate words from the documents in col. 5 lines 37-38.

It would have been obvious to one of ordinary skill in the art at the time the invention was made to have combined Kirsch, Kobayashi, and Turney into Kostoff to have created the claimed invention. It would have been obvious and desirable to have combined the punctuation and stop word removal technique of Kirsch into Kostoff so that the documents passes would

have been more efficient. It would have been obvious and desirable to have combined the synonym word replacement of Kobayashi into Kostoff so that the word counts could have been uniform across all of the documents, which would have yielded the most accurate clustering results. It would have been obvious and desirable to have combined the duplicate word removal of Turney into Kostoff so that the lists would have been uniform among all the documents in the cluster. This would have yielded the most accurate clustering results among the documents.

**Regarding dependent claim 3**, Kostoff teaches inputting one or more stop words, synonyms and a frequency level in col. 4 lines 39-49, col. 5 lines 59-64, and col. 6 lines 60-64.

**Regarding dependent claim 4**, Kostoff teaches adding words to a table in fig. 2, table 1, col. 4 lines 50-68, and col. 6 line 65 – col. 7 line 11. Kostoff teaches determining the frequency of each word remaining in the table in fig. 2, table 1, col. 4 lines 50-68, and col. 6 line 65 – col. 7 line 11. Kostoff teaches removing words below a frequency level from the table in col. 6 lines 2-64.

Kostoff does not teach removing punctuation and case from the documents. Kostoff does not teach removing stop words from the document. Kostoff does not teach replacing words in the documents with synonyms. Kostoff does not teach removing duplicate words from the documents. Kirsch teaches removing punctuation and case from the documents in col. 12 lines 5-7. Kirsch teaches removing stop words from the document in col. 12 lines 13-15. Kobayashi teaches replacing words in the documents with synonyms in fig. 3, 34-35, and col. 1 line 54 – col. 2 line 13. Turney teaches removing duplicate words from the documents in col. 5 lines 37-38.

It would have been obvious to one of ordinary skill in the art at the time the invention was made to have combined Kirsch, Kobayashi, and Turney into Kostoff to have created the claimed invention. It would have been obvious and desirable to have combined the punctuation and stop word removal technique of Kirsch into Kostoff so that the documents passes would have been more efficient. It would have been obvious and desirable to have combined the synonym word replacement of Kobayashi into Kostoff so that the word counts could have been uniform across all of the documents, which would have yielded the most accurate clustering results. It would have been obvious and desirable to have combined the duplicate word removal of Turney into Kostoff so that the lists would have been uniform among all the documents in the cluster. This would have yielded the most accurate clustering results among the documents.

**Regarding dependent claim 5**, Kostoff teaches inputting one or more stop words, synonyms and a frequency level in col. 4 lines 39-49, col. 5 lines 59-64, and col. 6 lines 60-64.

**Regarding dependent claim 7**, Kostoff teaches adding words to a dictionary table in fig. 2, table 1, col. 4 lines 50-68, and col. 6 line 65 – col. 7 line 11. Kostoff teaches determining the frequency of each word remaining in the table in fig. 2, table 1, col. 4 lines 50-68, and col. 6 line 65 – col. 7 line 11. Kostoff teaches removing words below a frequency level from the dictionary table in col. 6 lines 2-64.

Kostoff does not teach removing punctuation and case from the documents. Kostoff does not teach removing stop words from the document. Kostoff does not teach replacing words in the documents with synonyms. Kostoff does not teach removing duplicate words from the documents. Kirsch teaches removing punctuation and case from the documents in col. 12 lines 5-7. Kirsch teaches removing stop words from the document in col. 12 lines 13-15. Kobayashi

Art Unit: 2176

teaches replacing words in the documents with synonyms in fig. 3, 34-35, and col. 1 line 54 – col. 2 line 13. Turney teaches removing duplicate words from the documents in col. 5 lines 37-38.

It would have been obvious to one of ordinary skill in the art at the time the invention was made to have combined Kirsch, Kobayashi, and Turney into Kostoff to have created the claimed invention. It would have been obvious and desirable to have combined the punctuation and stop word removal technique of Kirsch into Kostoff so that the documents passes would have been more efficient. It would have been obvious and desirable to have combined the synonym word replacement of Kobayashi into Kostoff so that the word counts could have been uniform across all of the documents, which would have yielded the most accurate clustering results. It would have been obvious and desirable to have combined the duplicate word removal of Turney into Kostoff so that the lists would have been uniform among all the documents in the cluster. This would have yielded the most accurate clustering results among the documents.

**Regarding dependent claim 8**, Kostoff teaches inputting one or more stop words, synonyms and a frequency level in col. 4 lines 39-49, col. 5 lines 59-64, and col. 6 lines 60-64.

**Regarding dependent claim 9**, Kostoff teaches adding words to a table in fig. 2, table 1, col. 4 lines 50-68, and col. 6 line 65 – col. 7 line 11. Kostoff teaches determining the frequency of each word remaining in the table in fig. 2, table 1, col. 4 lines 50-68, and col. 6 line 65 – col. 7 line 11. Kostoff teaches removing words below a frequency level from the table in col. 6 lines 2-64.

Kostoff does not teach removing punctuation and case from the documents. Kostoff does not teach removing stop words from the document. Kostoff does not teach replacing words in

Art Unit: 2176

the documents with synonyms. Kostoff does not teach removing duplicate words from the documents. Kirsch teaches removing punctuation and case from the documents in col. 12 lines 5-7. Kirsch teaches removing stop words from the document in col. 12 lines 13-15. Kobayashi teaches replacing words in the documents with synonyms in fig. 3, 34-35, and col. 1 line 54 – col. 2 line 13. Turney teaches removing duplicate words from the documents in col. 5 lines 37-38.

It would have been obvious to one of ordinary skill in the art at the time the invention was made to have combined Kirsch, Kobayashi, and Turney into Kostoff to have created the claimed invention. It would have been obvious and desirable to have combined the punctuation and stop word removal technique of Kirsch into Kostoff so that the documents passes would have been more efficient. It would have been obvious and desirable to have combined the synonym word replacement of Kobayashi into Kostoff so that the word counts could have been uniform across all of the documents, which would have yielded the most accurate clustering results. It would have been obvious and desirable to have combined the duplicate word removal of Turney into Kostoff so that the lists would have been uniform among all the documents in the cluster. This would have yielded the most accurate clustering results among the documents.

**Regarding dependent claim 10**, Kostoff teaches inputting one or more stop words, synonyms and a frequency level in col. 4 lines 39-49, col. 5 lines 59-64, and col. 6 lines 60-64.

**Regarding dependent claim 12**, Kostoff teaches adding words to a dictionary table in fig. 2, table 1, col. 4 lines 50-68, and col. 6 line 65 – col. 7 line 11. Kostoff teaches determining the frequency of each word remaining in the table in fig. 2, table 1, col. 4 lines 50-68, and col. 6

Art Unit: 2176

line 65 – col. 7 line 11. Kostoff teaches removing words below a frequency level from the dictionary table in col. 6 lines 2-64.

Kostoff does not teach removing punctuation and case from the documents. Kostoff does not teach removing stop words from the document. Kostoff does not teach replacing words in the documents with synonyms. Kostoff does not teach removing duplicate words from the documents. Kirsch teaches removing punctuation and case from the documents in col. 12 lines 5-7. Kirsch teaches removing stop words from the document in col. 12 lines 13-15. Kobayashi teaches replacing words in the documents with synonyms in fig. 3, 34-35, and col. 1 line 54 – col. 2 line 13. Turney teaches removing duplicate words from the documents in col. 5 lines 37-38.

It would have been obvious to one of ordinary skill in the art at the time the invention was made to have combined Kirsch, Kobayashi, and Turney into Kostoff to have created the claimed invention. It would have been obvious and desirable to have combined the punctuation and stop word removal technique of Kirsch into Kostoff so that the documents passes would have been more efficient. It would have been obvious and desirable to have combined the synonym word replacement of Kobayashi into Kostoff so that the word counts could have been uniform across all of the documents, which would have yielded the most accurate clustering results. It would have been obvious and desirable to have combined the duplicate word removal of Turney into Kostoff so that the lists would have been uniform among all the documents in the cluster. This would have yielded the most accurate clustering results among the documents.

**Regarding dependent claim 13**, Kostoff teaches inputting one or more stop words, synonyms and a frequency level in col. 4 lines 39-49, col. 5 lines 59-64, and col. 6 lines 60-64.

**Regarding dependent claim 14**, Kostoff teaches adding words to a table in fig. 2, table 1, col. 4 lines 50-68, and col. 6 line 65 – col. 7 line 11. Kostoff teaches determining the frequency of each word remaining in the table in fig. 2, table 1, col. 4 lines 50-68, and col. 6 line 65 – col. 7 line 11. Kostoff teaches removing words below a frequency level from the table in col. 6 lines 2-64.

Kostoff does not teach removing punctuation and case from the documents. Kostoff does not teach removing stop words from the document. Kostoff does not teach replacing words in the documents with synonyms. Kostoff does not teach removing duplicate words from the documents. Kirsch teaches removing punctuation and case from the documents in col. 12 lines 5-7. Kirsch teaches removing stop words from the document in col. 12 lines 13-15. Kobayashi teaches replacing words in the documents with synonyms in fig. 3, 34-35, and col. 1 line 54 – col. 2 line 13. Turney teaches removing duplicate words from the documents in col. 5 lines 37-38.

It would have been obvious to one of ordinary skill in the art at the time the invention was made to have combined Kirsch, Kobayashi, and Turney into Kostoff to have created the claimed invention. It would have been obvious and desirable to have combined the punctuation and stop word removal technique of Kirsch into Kostoff so that the documents passes would have been more efficient. It would have been obvious and desirable to have combined the synonym word replacement of Kobayashi into Kostoff so that the word counts could have been uniform across all of the documents, which would have yielded the most accurate clustering results. It would have been obvious and desirable to have combined the duplicate word removal

of Turney into Kostoff so that the lists would have been uniform among all the documents in the cluster. This would have yielded the most accurate clustering results among the documents.

**Regarding dependent claim 15**, Kostoff teaches inputting stop words in col. 4 lines 39-49, col. 5 lines 59-64, and col. 6 lines 60-64.

**Regarding dependent claim 16**, Kostoff teaches inputting synonyms in col. 4 lines 39-49, col. 5 lines 59-64, and col. 6 lines 60-64.

**Regarding dependent claim 17**, Kostoff teaches inputting a frequency level in col. 4 lines 39-49, col. 5 lines 59-64, and col. 6 lines 60-64.

#### **(10) Response to Argument**

The Examiner believes Kostoff provides two particularly important teachings which would have suggested to one of ordinary skill in the art at the time of the invention to have created the claimed invention. One is that Kostoff does teach limiting the dictionary size to the preference of the user as is taught in fig. 2 and col. 5 line 1 – col. 6 line 64. Here Kostoff teaches that the user may select the highest frequency words and phrases and use those as the claimed dictionary to represent the pervasive themes of the document. Therefore, Kostoff teaches that it is useful and important for the user to select a portion of the document to use for its theme analysis. The second teaching the Examiner finds to be particularly pertinent to the claimed invention is the use of a trivial phrase list as taught by Kostoff in col. 4 lines 39-49 which may include any trivial phrases the user chooses to include. The number of trivial phrases that can be selected is unlimited except for limitations imposed by computer memory size. Therefore, the Examiner believes this teaching of Kostoff demonstrates that the text content of the document is



Art Unit: 2176

filtered prior to constructing the dictionary and thus the dictionary as taught by Kostoff contains less than all words in the document. Therefore, with this groundwork of teachings, the Examiner believes Kostoff provides teachings or suggestions for every limitation of the invention of claims 1, 6, and 11.

Regarding Appellant's argument in pages 11-16 that Kostoff does not teach or suggest all the limitations of independent claims 1 and 11, the Examiner respectfully disagrees. Kostoff teaches that the frequency of the words in a document is determined and the highest frequency words selected for a dictionary in fig. 2 and col. 5 lines 1 – col. 6 line 64. The number of words selected is determined by the user's preference or criteria set up by the user as taught in col. 6 lines 2-64. Therefore, Kostoff clearly teaches a motivation to truncate the dictionary to a maximum size selected by the user. Kostoff specifically teaches that the word and phrase selection is performed after the initial creation of the dictionary. However, Kostoff provides a teaching, which the Examiner believes is an important motivation, for filtering the text prior to the initial creation of the dictionary. In col. 4 lines 39-49, Kostoff teaches that a list of trivial phrases, created according to user selections, is excluded from the database prior to the creation of the dictionary. Therefore, Kostoff clearly teaches that the dictionary contains less than all of the words of the document because only the words considered to be nontrivial are used in the creation of the dictionary. Since Kostoff teaches a maximum number, or size, selection of highest frequency words, the Examiner believes the lower frequency words failing to meet the threshold can be considered to be trivial. The identification of these lower frequency trivial words is determined by the user when the user selects the number of high frequency terms to retain. Thus, since Kostoff teaches removing trivial words from the text prior to dictionary

Art Unit: 2176

creation and also teaches the identification of low frequency trivial words, the Examiner believes Kostoff teaches an obvious combination which reflects the maximum dictionary size limitation of the claimed invention. Under the view of this obvious combination, the claimed invention is taught because low frequency terms are identified as trivial and then the trivial terms determined to be not important by the user due to their low frequency counts are removed from the text prior to the creation of the dictionary. The dictionary will then only contain the highest frequency words and the phrases will be composed of only the highest frequency words. Thus, Kostoff under this self-suggested and self-motivated modification produces the same time and space efficient dictionary as the invention of claims 1 and 11.

Appellant claims in page 14 that Kostoff teaches using the entire full-text database to create lists of phrases and is therefore different from the claimed invention. However, the Examiner notes that Kostoff teaches in col. 4 lines 39-49 that the trivial list of phrases determined by the user is removed from the database prior to the creation of the dictionary and therefore the dictionary of Kostoff is really created from a filtered version of the full text of the document just as the claimed invention creates a dictionary from a filtered version of the full text of the document. Therefore, the Examiner respectfully disagrees with Appellant's position regarding the difference of the text used in creating Appellant's dictionary and the text used in creating the dictionary in Kostoff. With the frequency selection teaching of Kostoff, for example in fig. 2 and col. 5 line 1 – col. 6 line 64, the Examiner believes Kostoff suggests that the trivial words could be removed based on low frequency counts. In this view, the Examiner believes Kostoff suggests the use of a "maximum dictionary size" to limit word membership in the

Art Unit: 2176

generated dictionary to nontrivial high frequency words. Therefore, the Examiner respectfully maintains that the rejection of claims 1 and 11 should be sustained.

Regarding Appellant's argument in pages 16 and 17 that Kostoff does not teach or suggest all the limitations of independent claim 6, the Examiner respectfully disagrees. As shown in the remarks above with regard claims 1 and 11, the Examiner believes Kostoff suggests the use of a "maximum dictionary size" to limit word membership in the generated dictionary to nontrivial high frequency words. The claimed invention defines a first pass through the document which determines the most frequently occurring words. The claimed invention then defines a second pass through the document which determines the frequency of phrases in each document that contain only the high frequency words in the dictionary. Kostoff specifically teaches that the most frequently occurring words and most frequently occurring phrases are determined at the same time in one pass through the document, thereby not specifically teaching the two separate passes through the document. However, Kostoff does teach that the dictionary is composed in two steps. In the first step the trivial words are removed from the text database so that only the nontrivial words are used in the creation of the dictionary. In the second step of Kostoff, the frequency is determined for the words and phrases. Under the suggestion of Kostoff to have modified the trivial word removal step by removing the low frequency trivial words as described above, the Examiner believes this would have necessitated a two pass process as is defined in independent claim 6. Therefore, the Examiner believes Kostoff suggests creating the dictionary for the document in two passes with the first pass determining the frequency of the words in the document and removing the trivial low frequency words according to the user's selection of the number of high frequency nontrivial words to be included in the dictionary.

Art Unit: 2176

Kostoff, under this obvious modification, then performs the claimed second pass of determining the frequency of phrases in the document that are only composed of the high frequency words. Thus, the Examiner believes that the same obvious modification using the teachings and suggestions of Kostoff for teaching a “maximum dictionary size” also necessarily teaches two passes through the document. Therefore, the Examiner respectfully maintains that the rejection of claim 6 should be sustained.

Regarding Appellant’s argument in pages 25 and 26 that Kostoff, Kirsch, Kobayashi, and Turney do not teach or suggest all the limitations of dependent claims 2-5, 7-10, and 12-17, the Examiner respectfully disagrees. The Examiner reasserts in these remarks that Kirsch, Kobayashi, and Turney teach replacing words in a document with synonyms, removing punctuation and case, removing stop words, and removing duplicate words from a document. Kostoff teaches in col. 4 lines 39-49 that the user may remove any trivial words from the text database before creation of the dictionary. Kostoff teaches limiting the dictionary size to the preference of the user in accordance with the frequency of the words in the document in fig. 2 and col. 5 line 1 – col. 6 line 64. The Examiner believes the combination of these teachings and suggestions enables Kostoff to remove more than just stop words as Appellant asserts in page 25. Therefore, the Examiner maintains that Kostoff does not teach away from the claimed “maximum dictionary size” used to limit the dictionary. Kostoff teaches the words removed from the full-text database may be determined by the user and thus a word which is trivial in the creation of one dictionary, could be nontrivial in the generation another dictionary. Therefore, the Examiner believes that the combination of Kostoff, Kirsch, Kobayashi, and Turney teach or

Art Unit: 2176

suggest all the limitations of dependent claims 2-5, 7-10, and 12-17. The Examiner respectfully maintains that the rejection of claims 2-5, 7-10, and 12-17 should be sustained.

**(11) Related Proceeding(s) Appendix**

No decision rendered by a court or the Board is identified by the examiner in the Related Appeals and Interferences section of this examiner's answer.

For the above reasons, it is believed that the rejections should be sustained.

Respectfully submitted,

Peter J. Smith *Peter J. Smith* 1/12/2006

Conferees:

Heather Herndon

*Heather Herndon*  
HEATHER R. HERNDON  
SUPERVISORY PATENT EXAMINER  
TECHNOLOGY CENTER 2100

William Bashore

*William L. Bashore*  
WILLIAM BASHORE  
PRIMARY EXAMINER  
1/17/2006

Stephen Hong

*Stephen Hong*  
STEPHEN HONG  
SUPERVISORY PATENT EXAMINER